NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

THE EFFECT OF GENDER ON ATTRITION AT THE DEFENSE LANGUAGE INSTITUTE FOREIGN LANGUAGE CENTER

by

George T. Arthur

September, 1996

Thesis Advisor: Lyn R. Whitaker

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE September 1996	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE
The Effect of Gender on Attrition at the Defense Language Institute Foreign Language Center.

5. FUNDING NUMBERS

6. AUTHOR(S) Arthur, George T.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

The Defense Language Institute Foreign Language Center (DLIFLC) , located at the Presidio of Monterey, California, provides language training for Department of Defense military and civilian personnel. The Institute trains approximately 2,500 students annually, of which approximately 26 percent are female. Student attrition is a costly feature of this training program. Females experience roughly a 7 percent higher rate of attrition than males at DLIFLC. The Institute is interested in knowing whether this difference indicates a gender bias, or whether it can be explained by other factors. This study investigates this question. Specifically, data on FY-95 DLIFLC students are examined to determine factors with a significant impact on attrition, with particular emphasis on gender. Such information is potentially useful to the Institute for internal quality assurance efforts as well as part of potential cost saving measures.

14. SUBJECT TERMS
Gender, Attrition, Language Training, DLI

15. NUMBER OF PAGES
58

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

Approved for public release; distribution is unlimited.

THE EFFECT OF GENDER ON ATTRITION AT THE DEFENSE LANGUAGE INSTITUTE FOREIGN LANGUAGE CENTER

George T. Arthur
Lieutenant, United States Navy
B.S., United States Naval Academy, 1986

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1996

Author: ____________________
George T. Arthur

Approved by: ____________________
Lyn R. Whitaker, Thesis Advisor

____
Stephen M. Payne, Second Reader

____
Frank C. Petho, Chairman
Department of Operations Research

ABSTRACT

The Defense Language Institute Foreign Language Center (DLIFLC), located at the Presidio of Monterey, California, provides language training for Department of Defense military and civilian personnel. The Institute trains approximately 2,500 students annually, of which approximately 26 percent are female. Student attrition is a costly feature of this training program. Females experience roughly a 7 percent higher rate of attrition than males at DLIFLC. The Institute is interested in knowing whether this difference indicates a gender bias, or whether it can be explained by other factors. This study investigates this question. Specifically, data on FY-95 DLIFLC students are examined to determine factors which have a significant impact on attrition, with particular emphasis on gender. Such information is useful to the Institute for internal quality assurance efforts as well as part of potential cost saving measures.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Defense Language Institute Foreign Language Center (DLIFLC), located at the Presidio of Monterey, California, provides language training for Department of Defense military and civilian personnel. The Institute trains approximately 2,500 students annually, of which approximately 26 percent are female. Student attrition is a costly feature of this training program. Females experience roughly a 7 percent higher rate of attrition than males at DLIFLC. The Institute has asked whether this difference is an indication of potential gender bias, or is it a function of other characteristics? This study investigates this question.

The methodology used for this study involves fitting a logistic regression model with graduation/attrition as the response, and a variety of demographic, language specific, and test score variables as predictors. By analyzing variables with a significant effect on the model, it is possible to identify factors which contribute to student attrition, with particular emphasis on gender.

Data are obtained from the combined Defense Language Institute Foreign Language Center - Defense Manpower Data Center data base, and include students scheduled to graduate in FY-95. There are 1,985 students in the data used for this study.

Separate models are run on aggregate data and on individual service groups. For the aggregate data, the interaction between gender and service branch is a significant predictor of attrition. This is because, for Air Force students, gender itself is a significant predictor of attrition. Other attributes are different for Air Force students as well. The proportion of females for Air Force students is higher than for the other services. Also, a higher percentage of Air Force females are in the more difficult (Category IV) languages at DLIFLC. Finally, Air Force females are mostly in paygrades E-3 and below; students in these paygrades tend to be at a higher risk for attrition. Preliminary results show that the higher attrition statistics for females are not likely due to their gender; rather, females are over-represented in certain 'high risk' groups.

In general, for all students, *language difficulty category* and *prior language experience* tend to have the most impact on attrition, followed by certain demographic variables and test scores. Further study is suggested on the issues concerning Air Force students, and on the specific reasons why students fail to graduate (i.e., academic, administrative, etc.).

The information gained from this study should assist the Institute with internal quality assurance measures, and provide it with a better understanding of the relationship between gender and attrition at DLIFLC.

I. INTRODUCTION

The Defense Language Institute Foreign Language Center (DLIFLC) is located at the United States Army Presidio of Monterey, California. The Institute is responsible for training military members from all four service branches, as well as civilian Federal employees, in a variety of missions requiring knowledge of a foreign language. The Institute produces approximately 2,500 graduates annually. (Directorate for Academic Administration, 1995)

A. PROBLEM STATEMENT

At the DLIFLC approximately 26 percent of the student population are female. DLIFLC FY-95 data indicate that the attrition rate among females is approximately 34 percent, while that of males is approximately 27 percent (Figure 1). By comparison, FY-95 Army-wide attrition for Initial Entry Trainees[1] (IETs) is approximately 16 percent among females and 10 percent among males (Dove, 1996)[2]. Does the 7% difference in overall attrition for DLIFLC students indicate the existence of gender bias or is the difference a manifestation of other factors (e.g., a higher percentage of female students in more difficult curricula or a function of general differences in attrition among IETs in general)? Interest in gender-related attrition at DLIFLC goes back at least two decades; a 1975 point paper entitled Army Linguist Personnel Study (ALPS) cited attrition statistics which were remarkably similar to contemporary numbers, with overall female attrition of 34.6%, and overall male attrition of 27% (Rice, 1975). The Institute is interested in further exploration of these issues, and this study does so. The information provided by this study will assist the Institute with internal quality

[1] Initial Entry Trainees are those soldiers who have not yet completed their Basic and Advanced individual training.

[2] This study does not address the difference between DLIFLC attrition statistics and those of IETs in other training programs. Its focus is on attrition within DLIFLC.

assurance efforts, as well as provide potentially useful information to the chain of command.

While there is little background literature addressing the unique environment of military language training, the effect of gender on first language development is relatively well-documented. In general, females learn to talk and use sentences earlier than males, and are shown to use a greater variety of words (O'Mara, 1994). Furthermore, from about the sixth grade through college, females consistently outscore males on a variety of measures of verbal skills (O'Mara, 1994). The exact reason for these differences is unknown. Neurological studies have shown, however, that there are physiological differences between the brains of males and females. These differences include the presence of more neurons and increased size in areas of the brain associated with language function. These physiological differences as well as the effects of differing cultural expectations are thought to be significant. (Begley, 1995)

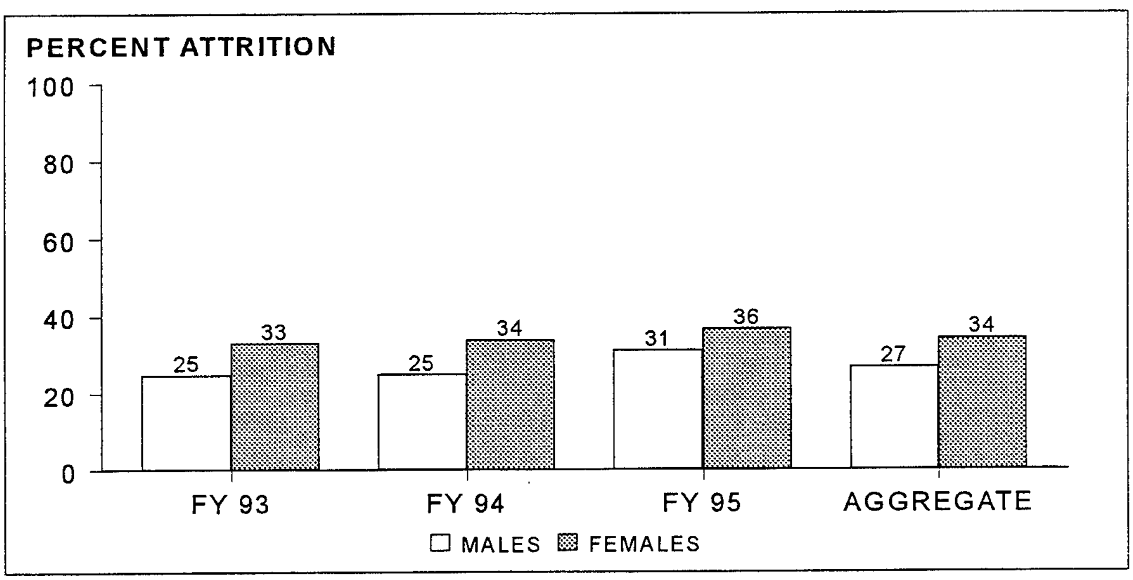

Figure 1. Percentage of male/female students who failed to graduate with their class.

It is reasonable to assume that this advantage in aptitude among females would manifest itself in second language learning as well. This is an apparent contradiction to the attrition statistics shown in Figure 1. It is interesting to note that although *overall* attrition among females is higher than for males, *academic* attrition among females is approximately 15% lower than for males[3] (Figure 2). The 1975 ALPS study found a 9% lower academic attrition rate for females. This comparison suggests a possible explanation for the contradiction; i.e., it is possible that factors unrelated to academic performance

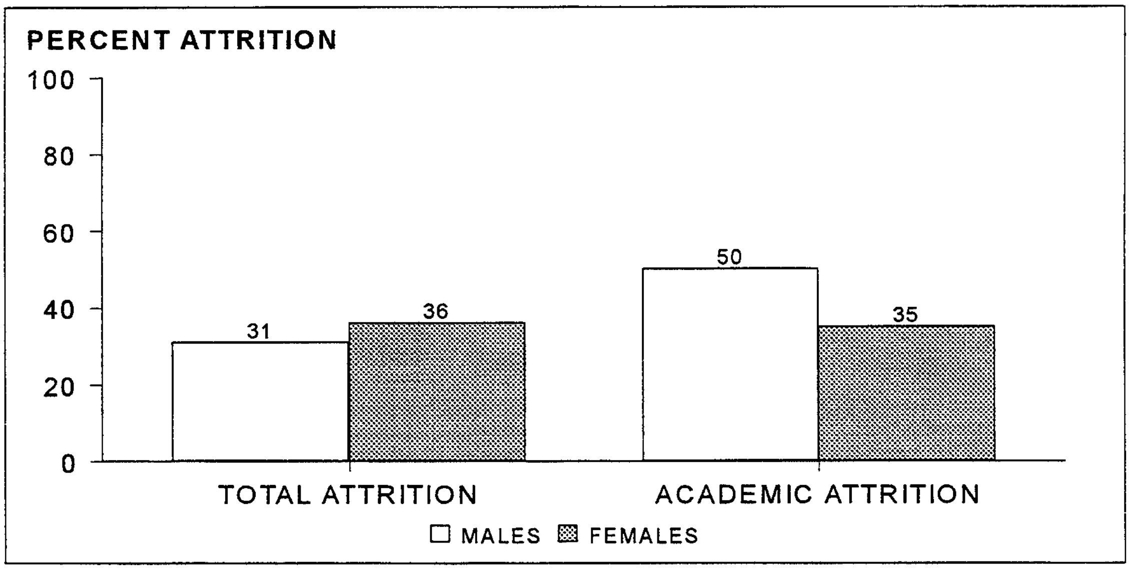

Figure 2. Students scheduled to graduate in FY 95. Comparison of total attrition vs. percentage of non-graduates who attrited for academic reasons.

may account for the higher overall attrition rate among females. This issue will be explored further in this study.

B. LANGUAGE SKILL CHANGE PROJECT

The results of a study similar to this thesis were released in August of 1994. The study, entitled Language Skill Change Project (LSCP), was

---

[3] Overall attrition refers to students who fail to graduate for any reason. Academic attrition refers to students who fail to graduate specifically due to academic performance.

conducted by the DLIFLC Research and Analysis Division with the support of PRC, Inc., a civilian contractor. The LSCP reported no specific conclusions about the effect of gender on attrition, although gender was a sub-factor in a predictor block including various demographic variables. The predictor block including sex, level of education, and age was found to be collectively significant. (O'Mara, 1994)

There are several key areas in which this study differs from the LSCP study. The first is scope. The main focus of the LSCP was to track changes in language proficiency (listening, reading, and speaking) over time. While language training attrition was addressed in the LSCP, it was not the primary emphasis, and was restricted to academic attrition (O'Mara, 1994). This study addresses language training attrition of all types, and language proficiency is not addressed.

The second area in which the two studies differ is in the subject population. The LSCP included only U.S. Army personnel who had, or were preparing for, military intelligence linguist occupational specialty codes, who were enrolled in either Spanish, German, Russian or Korean (one language in each of the four language difficulty categories). This study includes students in all branches of the military, and spans all applicable languages and language difficulty categories. While the LSCP was a longitudinal study, tracking students' progress over a 3 to 4 year period, this study is a cross-sectional study, including those students who were scheduled to graduate during FY-95, and includes 1,985 subjects.[4]

The third major area in which the studies differ is in the data. Data used in the LSCP included information available in the subjects' records, as well as a

---

[4] At the request of the Institute, the focus is on recent trends. FY 95 enrollees are chosen as this is the latest year for which complete data are available. Students who were 're-cycled' from prior classes in the same language or who were transferred from other languages are excluded. Re-cycling is the process of removing a student from his/her current class, and starting them over in a later class in the same language. This can occur for many reasons, such as poor academic performance, medical problems, etc.

series of special instruments used in assessing a variety of aptitudes, attitudes, motivational factors and personality-related characteristics (O'Mara, 1994). Data used in this study includes information available from current records, and does not incorporate any special testing instruments or surveys not normally administered to the language trainee population as a whole.

C. THESIS ORGANIZATION

Chapter II gives an overview of the data used to conduct this study. It contains an explanation of the data source, and the methods used to identify relevant variables. Variables selected for use in modeling are explained in detail. Chapter III contains the bulk of the analysis. Preliminary data exploration is conducted on the variables selected in Chapter II. An explanation of the logistic regression model used in this study and its results are provided in Chapter III. Chapter IV summarizes final results, and provides conclusions and recommendations for further study.

II. DATA

The data gathered for this study are used in two stages. First, preliminary analysis is performed on each variable to determine which variables are suitable for inclusion as potential predictors of attrition. Second, variables identified in the first stage for inclusion are used to construct a regression model of attrition. Of particular interest is whether gender is a significant predictor of attrition. The preliminary analysis and selection process are discussed in this chapter, and further analysis stemming from the regression model is found in Chapter III.

A. THE SOURCE

This study is being conducted with the cooperation of the DLIFLC Research and Analysis Division and the Command Historian. Data are drawn from the combined DLIFLC - Defense Manpower Data Center (DMDC) Student Database (S3D). S3D represents a comprehensive aggregation of data elements extracted from DLIFLC's Student Data Base and DMDC's Active, Loss, Reserve and Civilian files. These files are large, containing thousands of records (one per individual) with over 350 data fields per record. concatenated by the students' social security number and updated quarterly. (Shaw, et al, 1994)

B. THE DATA

At the request of the Institute, emphasis is placed on recent trends. This is done to capture the effects of contemporary policies at the Institute, without consideration of changing effects over time. Therefore, this study concentrates on students who were scheduled to graduate during FY 95 because this is the most recent full year for which data are available. Students eligible for consideration are those considered as new inputs. This criterion eliminates students who were in intermediate or advanced classes, as well as those who were transferred from other languages or re-cycled from earlier classes in the same language. The rationale for this criterion is two-fold: 1) the excluded subjects are not considered typical of the student population at large, and therefore could introduce confounding effects in the analysis, and 2) the excluded subjects represent less than 10 percent of the target population and therefore do not constitute a significant portion of the population. All students who meet the above criteria are included in the data, resulting in 1,985 observations. The data includes students from each of the four language difficulty categories, and spans all four branches of the service.

C. VARIABLES

Each record in the database has 352 variables. Through in-depth consultation with subject matter experts at the Institute, 43 of these variables are identified as potential candidates for inclusion, and are defined in Table 1. Redundant variables are excluded, as well as those which clearly have no relevance to the question of attrition.

To simplify the modeling effort, it is necessary to further refine the set of candidate predictor variables. For each variable, the decision is to either exclude it, use it in its current form, or use it as a basis for some new transformed variable.

The binary response variable indicating graduation or attrition (GRAD/ATTR) is constructed from the variables output status (OUT) and reason for output (REASON). This is done by evaluating the output status and reason codes and determining whether a particular student successfully completed his/her curriculum on time. If so, they are labeled a graduate, otherwise they are placed into the attrition category.

The explanatory variables fall loosely into three categories: 1) demographic variables, 2) variables associated with the language studied at DLIFLC or prior language experience, and 3) variables associated with test results measuring learning aptitude or demonstrated ability.

VARIABLE	DESCRIPTION	DATA TYPE	NUMBER OF LEVELS
OUT	student output category	nominal	7
REASON	reason for in or out of class	nominal	37
SSN	social security number	nominal	N/A
SEX	gender	nominal	2
PAYGRD	paygrade	nominal	20
YRSRV	years of military service	continuous	N/A
EDUYR	years of education	continuous	N/A
MARRY	marital status	nominal	2
MOTIV	language choice - motivation	ordinal	5
DOB	date of birth	nominal	N/A
SERV	service	nominal	5
ETHNIC	race, ethnic	nominal	7
LID	language identification code	nominal	22
LENGTH	length of course (weeks)	nominal	N/A
PRILANG	prior language code	nominal	46
NATENG	native of english language	nominal	2
OTHER	native of other language	nominal	2
PRPROF	proficiency of prior language	ordinal	5
PRSRC	source of prior language	nominal	7
PREXP	experience of prior language	ordinal	8
LANCAT	language category	ordinal	4
GPA	grade point average (dliflc)	continuous	N/A
DLPTL	Defense Language Proficiency Test score (listening)	continuous	N/A
DLPTR	Defense Language Proficiency Test score (reading)	continuous	N/A
DLPTS	Defense Language Proficiency Test score(speaking)	continuous	N/A
DLAB	Defense Language Aptitude Battery Test score	continuous	N/A
AFQT	Armed Forces Qualification Test score	continuous	N/A
TESTV	Armed Forces Qualification Test form version	nominal	N/A
ASVFM	Armed Services Vocational Aptitude Battery test form version	nominal	N/A
GS	Armed Services Vocational Aptitude Battery test - general science	continuous	N/A
AR	Armed Services Vocational Aptitude Battery test - arithmetic reasoning	continuous	N/A
WK	Armed Services Vocational Aptitude Battery test - word knowledge	continuous	N/A

VARIABLE	DESCRIPTION	DATA TYPE	NUMBER OF LEVELS
PC	Armed Services Vocational Aptitude Battery test - paragraph comprehension	continuous	N/A
NO	Armed Services Vocational Aptitude Battery test - numeric operation	continuous	N/A
CS	Armed Services Vocational Aptitude Battery test - coding speed	continuous	N/A
AS	Armed Services Vocational Aptitude Battery test - auto and shop information	continuous	N/A
MK	Armed Services Vocational Aptitude Battery test - mathematics knowledge	continuous	N/A
MC	Armed Services Vocational Aptitude Battery test - mechanical comprehension	continuous	N/A
EI	Armed Services Vocational Aptitude Battery test - electronics information	continuous	N/A

Table 1. Variables downloaded from data base.

1. Demographic Variables

The following variables are related to demographics: gender, social security number, paygrade, years of service, years of prior education, marital status, motivation, age, branch of service, and ethnic background. The binary predictor variable describing gender (SEX) is included because this is the primary predictor of interest. As shown in Chapter I, Figure 1, there appears to be increased attrition among female students. The nominal variable listing a student's social security number (SSN) is excluded as this information is used for data management and has no impact on attrition.

The categorical variable indicating an observation's military paygrade (PAYGRD) contains 20 levels. Some of these levels have very few observations. For example, W-5 has only one observation. PAYGRD is therefore transformed into a continuous variable (PAYGRD2) in the following manner: each level of PAYGRD (E-1 through O-6) is arranged in increasing order, then is coded numerically. E-1 is assigned as '1', E-2 as '2' and so forth ending with O-6 assigned as '20'. PAYGRD is used as the basis for another categorical variable, indicating whether an observation is an officer or is enlisted (OFF/ENL). This variable contains two levels and is formed by assigning all observations with paygrade E-9 and below to the enlisted category and all others to the officer category. This variable is designed to detect any possible differences between officers and enlisted students with respect to attrition. PAYGRD2 and OFF/ENL are included in the data set. There appears to be a decreasing and then increasing rate of attrition among enlisted students as they become more senior in paygrade. A similar relationship exists among commissioned officers. There is no clear trend among warrant officers. The relationship between paygrade and attrition is depicted in Figure 3. It is interesting to note that a vast majority of students come from lower (E3 and below) paygrades (Figure 4.).

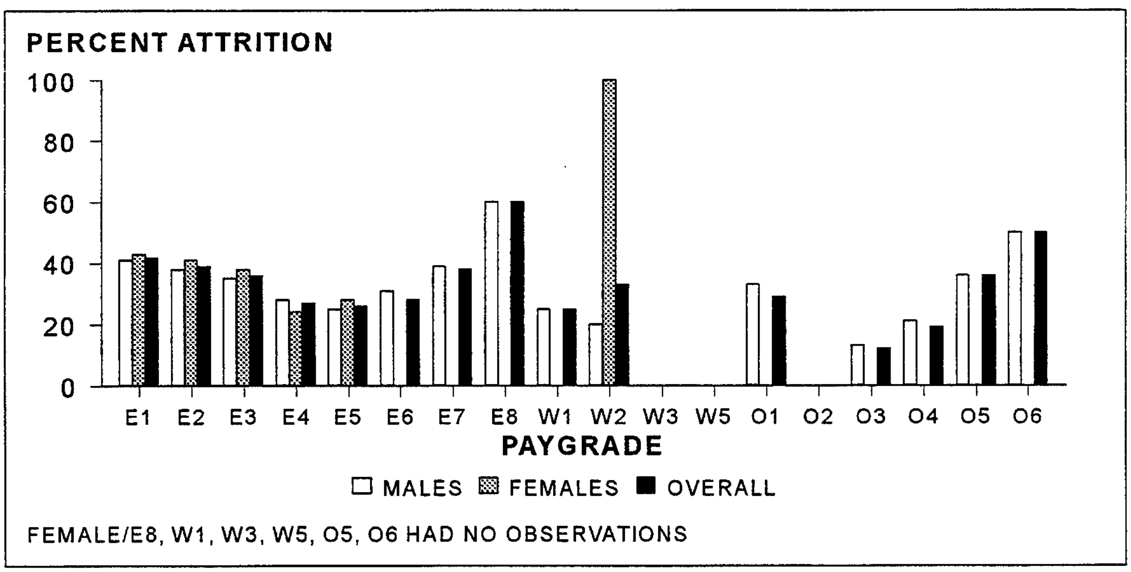

Figure 3. Percentage of attrition vs. paygrade.

Several of the predictor variables provide age type information. One of them is the continuous variable indicating years of military service (YRSRV). Although YRSRV may be redundant with PAYGRADE or other such variables, they are included in the study. In the case of YRSRV, the majority of observations (67%) have less than two years of service. For graphical purposes the observations are separated into those with fewer than two years of service

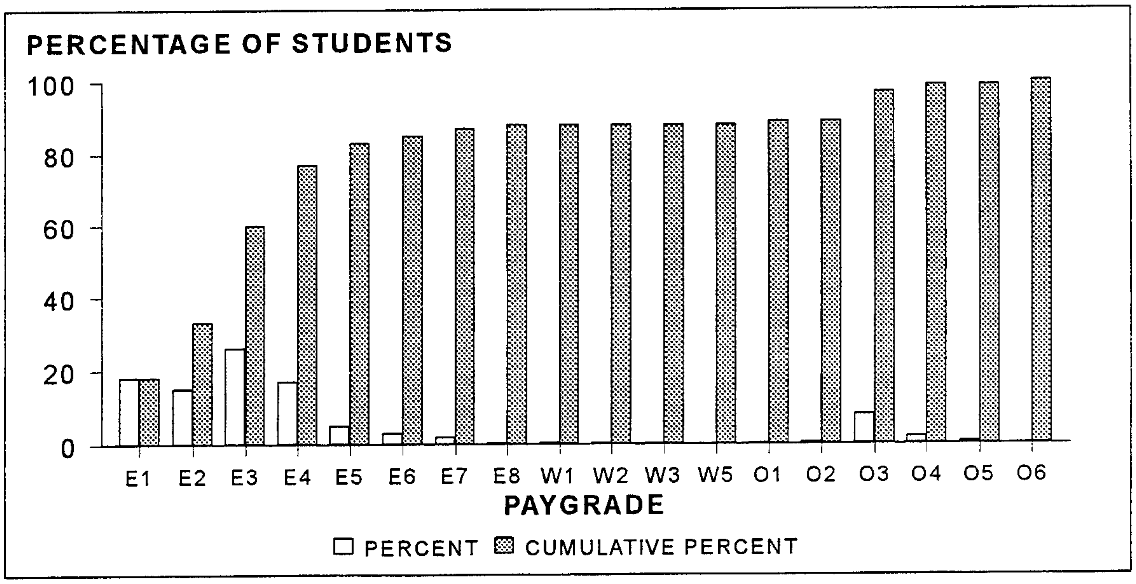

Figure 4. Paygrade distribution of subject data.

and those with two or more years of service. There appears to be a higher rate of attrition among observations in the former group, as depicted in Figure 5.

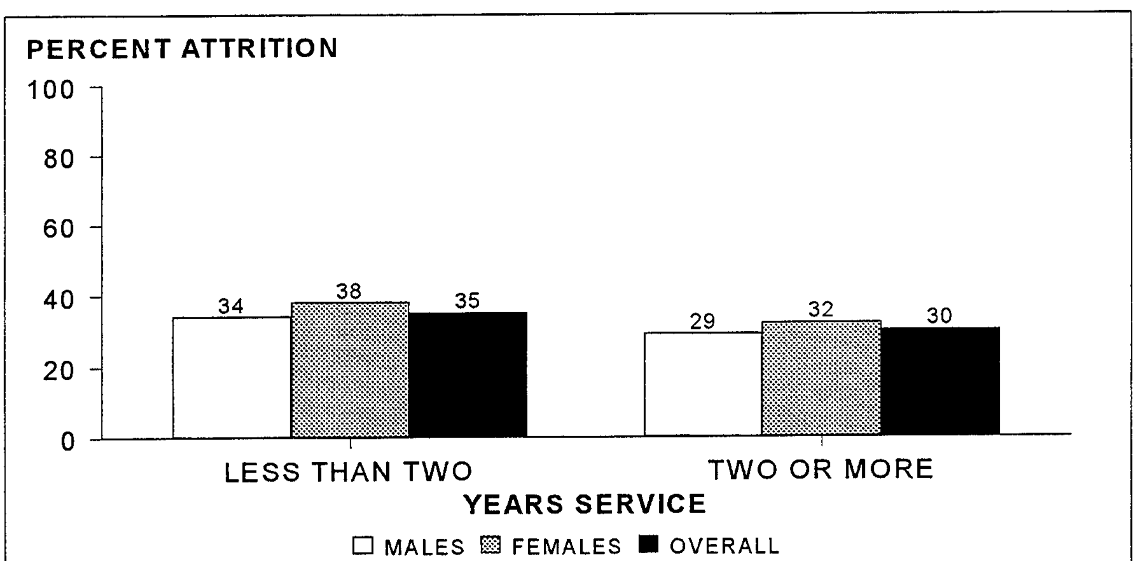

Figure 5. Percentage of attrition vs. years of military service.

In the data set used for this study, there are some occurrences of missing values. In the case of the continuous variable indicating years of education (EDUYR), approximately 20 percent of the observations have missing values. A common attribute of the missing data for this variable is that they are all attrites. There is no clear reason for this; it would be useful for future research purposes to determine the cause of this situation, and correct the data collection procedures, if necessary. Care needs to be exercised in the handling of missing values. If an observation has a missing value for any of its variables, that observation is usually excluded from analysis. To prevent the complete exclusion of observations with missing values for EDUYR, this variable is transformed from continuous to nominal. A new variable, EDUYRgroup, is formed by including all observations with missing values in one level (N/A), all observations with no more than a high school education in another level (HS), and all observations with some college in a third level (HS+). Thus, EDUYRgroup is included in the data set to detect possible effects of quantity of prior education on attrition. From Figure 6, students with some college have a lower percentage of attrition.

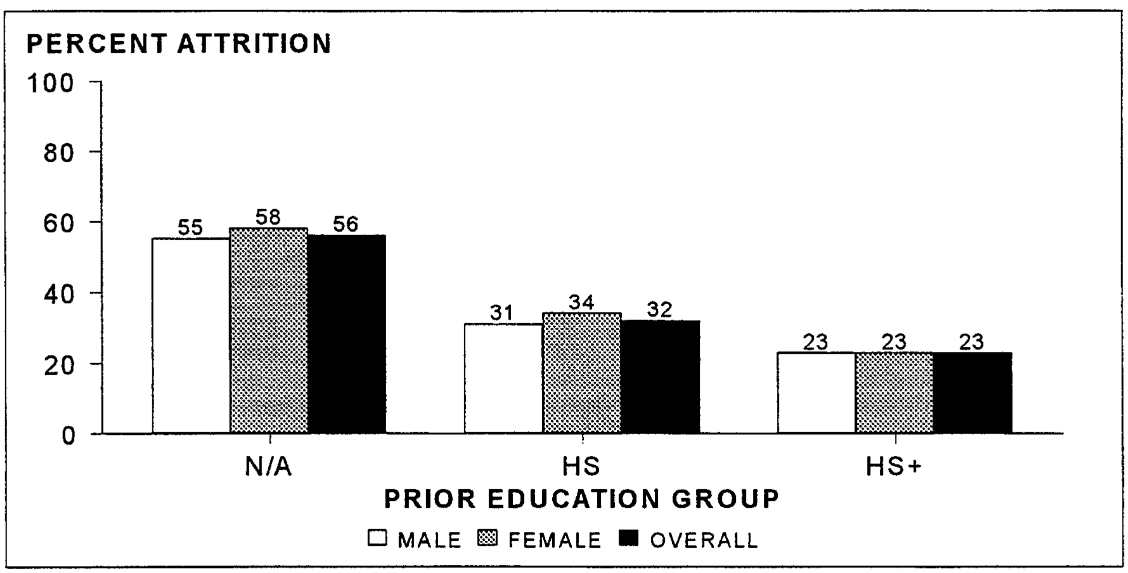

Figure 6. Percentage of attrition vs. prior education.

The binary variable indicating marital status (MARRY) is included to explore the possible effects of marital status on attrition. Overall, married students seem to have a lower percentage of attrition than single students. However, married females appear to experience a higher percentage of attrition than single females. Figure 7 shows the relationship between marital status and attrition.

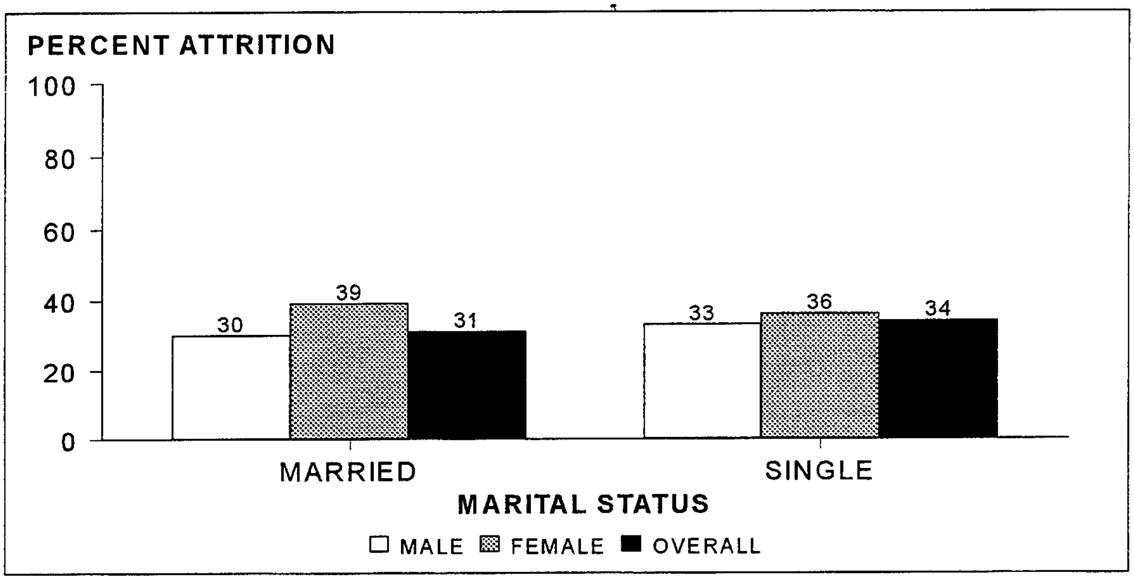

Figure 7. Percentage of attrition vs. marital status.

The ordinal variable describing a student's motivation to study the assigned language (MOTIV) contains 5 levels. They are self-evaluated by the student, and range from 1 (least motivated) to 5 (most motivated). This variable is included to examine the effects of motivation on attrition. From Figure 8, after level 2, there is a steady decline in percentage of attrition as motivation increases.

The variable indicating date of birth (DOB) was transformed into the variable AGE by computing a subject's age as of 01JAN95. AGE is included in the predictor set. For graphical purposes, AGE is broken into four age groups. From Figure 9 it appears that the percentage of attrition generally decreases with age. Thus AGE is treated as a continuous variable rather than as a categorical variable.

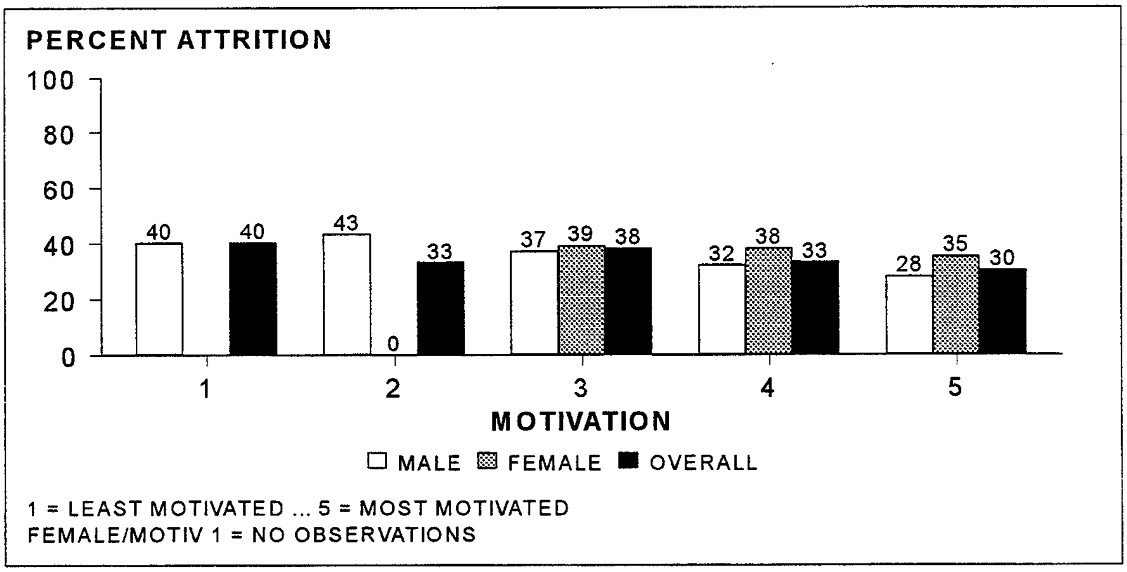

Figure 8. Percentage of attrition vs. motivation.

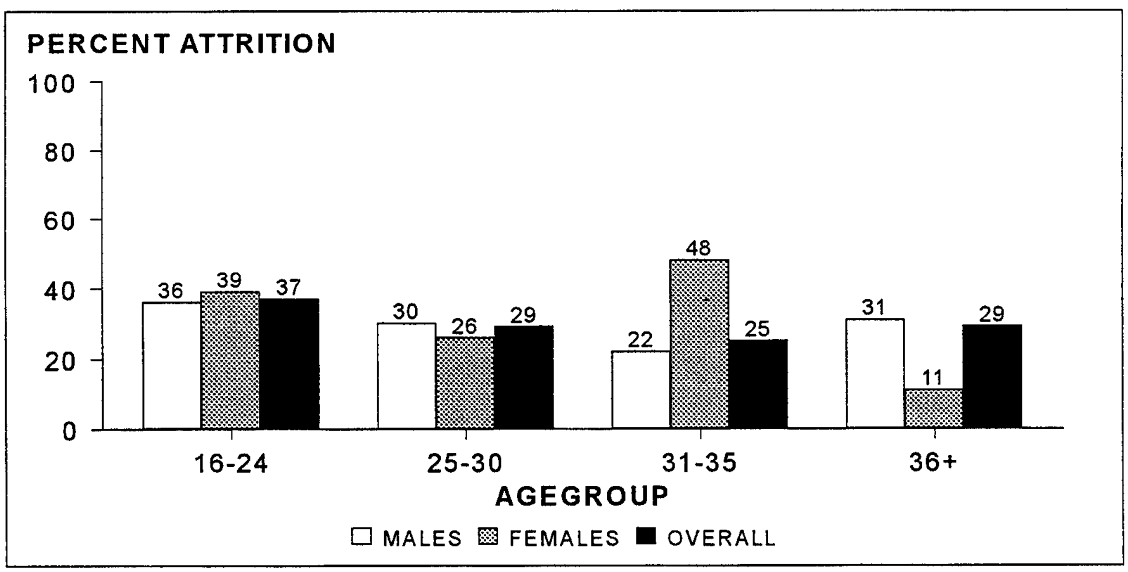

Figure 9. Percentage of attrition vs. age.

The categorical variable indicating which branch of service a student was in (SERV) is included to pick up any relationship between service component and attrition. From Figure 10, Army students had the highest overall attrition (36%) while Navy students had the lowest overall attrition (23%). The fact that female Marines experienced 60% attrition is potentially significant.

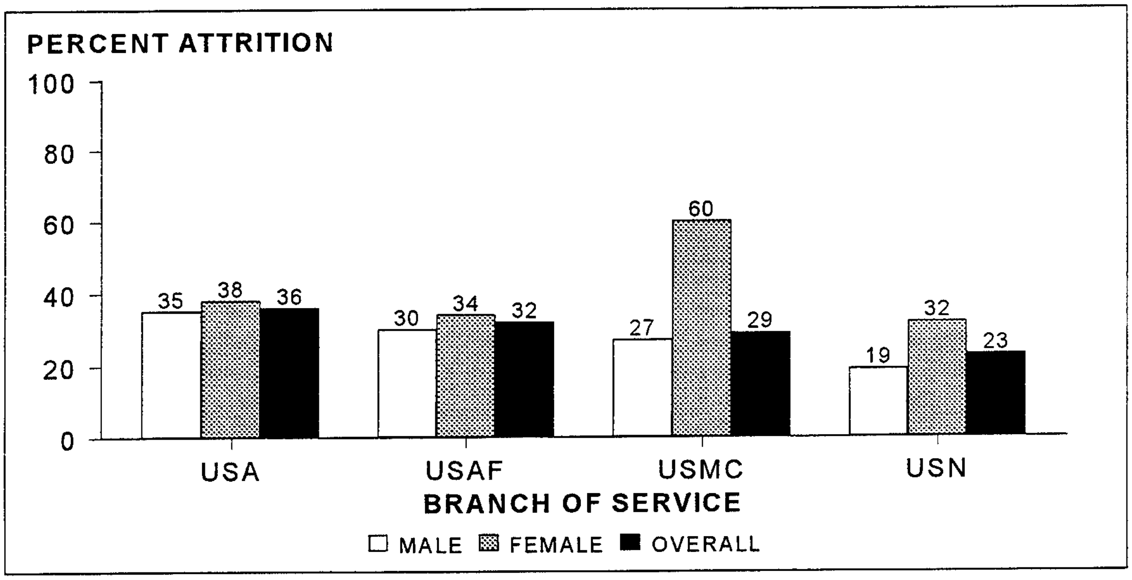

Figure 10. Percentage of attrition vs. branch of service.

The categorical variable describing a student's ethnic group (ETHNIC) is included to determine any effects of ethnic background on attrition. From Figure 11, there is wide variation in attrition percentage across different groups, ranging from a high of 57% overall attrition for those observations listed as 'unknown/none', to a low of 17% overall attrition for Hispanics.

2. Language Related Variables

The following variables are related to a student's language training and experience, both prior to and at DLIFLC: language category, language identification code, course length, prior language category, prior language experience level, prior language source, prior language proficiency, and whether a student is a native English speaker or of some other language.

The ordinal categorical variable indicating a student's language category (LANCAT) has four levels: I, II, III, IV. These levels indicate, in increasing order, the relative difficulty of a student's particular language curriculum in accordance with established guidelines at the Institute. This variable is included to show

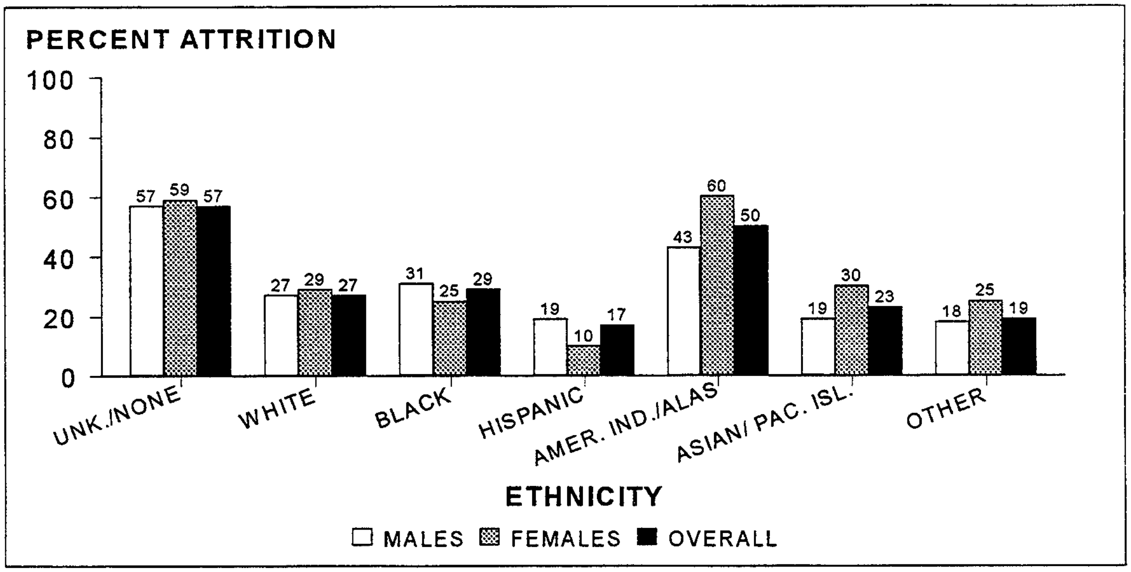

Figure 11. Percentage of attrition vs. ethnic group.

the effects of language difficulty on attrition. As shown in Figure 12, the two most difficult levels have a greater percentage of attrition.

The nominal variable indicating a student's language identification code (LID) specifies a unique code for each particular language curriculum. For the data in this study, this variable has 22 levels, some with too few observations to be useful. For example, Greek has only 3 observations. Since LID is a subset of LANCAT, and LANCAT contains the desired information (i.e., relative difficulty) LID is excluded in favor of LANCAT. The advantage of using LANCAT instead of LID is that it allows for the pooling of LID categories with relatively few observations into their respective language categories. The variable indicating a student's curriculum length in weeks (LENGTH) is excluded. This is because LENGTH varies as a function of language difficulty, and therefore the information provided by LENGTH is reflected in LANCAT.

The nominal variable indicating prior language experience is called prior language code (PRILANG). This variable is coded the same as LID, and for this data set has 46 levels. This variable is used as the basis for another variable,

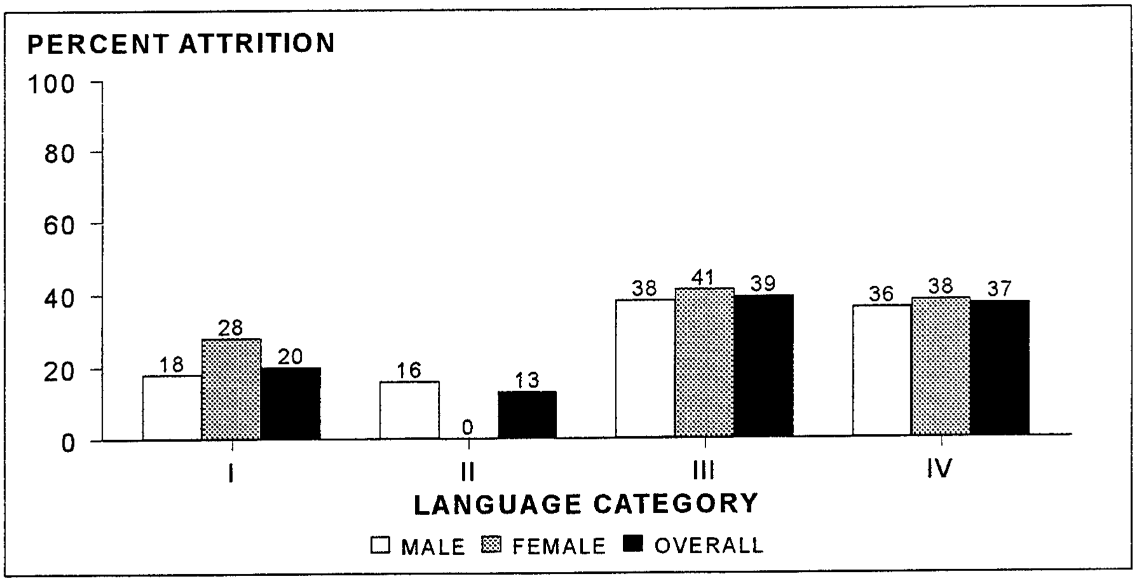

Figure 12. Percentage of attrition vs. language category.

prior language category (PRILANCAT). PRILANCAT is computed in the exact manner as LANCAT, by assigning each observation with prior language experience to its associated relative difficulty category. The variable PRILANCAT is included in favor of PRILANG for the same reasons that LANCAT is preferred over LID. An additional benefit of including PRILANCAT is that it is directly comparable to LANCAT. From Figure 13, students with no prior language experience have higher probabilities of attrition, second only to students with prior experience in category IV languages. Of students with prior language experience, there is an increased percentage of attrition among PRILANCAT IV students. Nominal variables indicating prior language experience level, prior language source, prior language proficiency, and whether a student is a native English speaker or of some other language (PREXP, PRSRC, PRPROF, NATENG, OTHER) are excluded. This is done because the desired information (i.e., relative difficulty of prior language, if any) is contained in the variable PRILANCAT.

3. Test Score Variables

The following variables are related to aptitude or performance measures: Armed Forces Vocational Aptitude Battery, Armed Forces Qualification Test, Defense Language Aptitude Battery, Defense Language Proficiency Tests, test form versions, and grade point average. The first three, Armed Services Vocational Aptitude Battery (ASVAB), Armed Forces Qualification Test (AFQT),

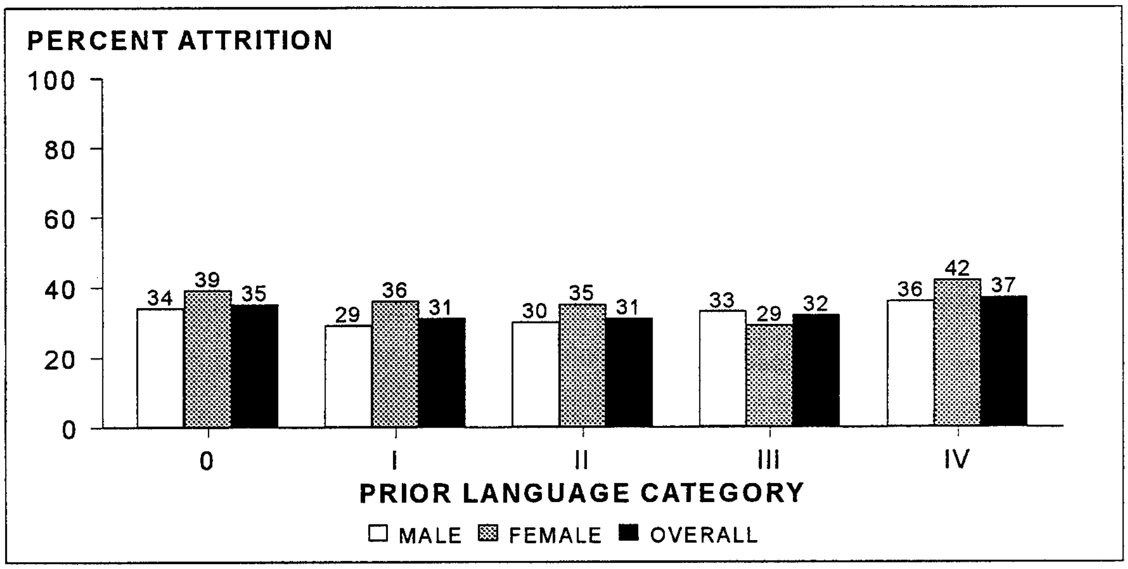

Figure 13. Percentage of attrition vs. prior language category.

and Defense Language Aptitude Battery (DLAB) are important to this study. The ASVABs are a battery of 10 tests administered to potential recruits measuring such skills as general science, paragraph comprehension, and mathematics knowledge. A complete listing of these sub tests is located in Table 1. The AFQTs are a composite measure formed from the ASVABs. The DLAB test is a specific measure of language learning aptitude, administered to language training candidates. The continuous variable DLAB is included to capture the effects of language learning aptitude on attrition. From Figure 14, there is a generally decreasing percentage of attrition as DLAB scores increase.

Many of the ASVAB sub tests measure similar types of aptitude. This redundancy in the tests can result in multicollinearity of the test scores. To guard against multicollinearity, and to potentially reduce the number of predictor variables, the method of principle components is used. Principle components is a technique that results in orthogonal linear combinations of the predictor variables (or standardized versions of the predictor variables). The first principle component is the linear combination of the predictor variables that has the

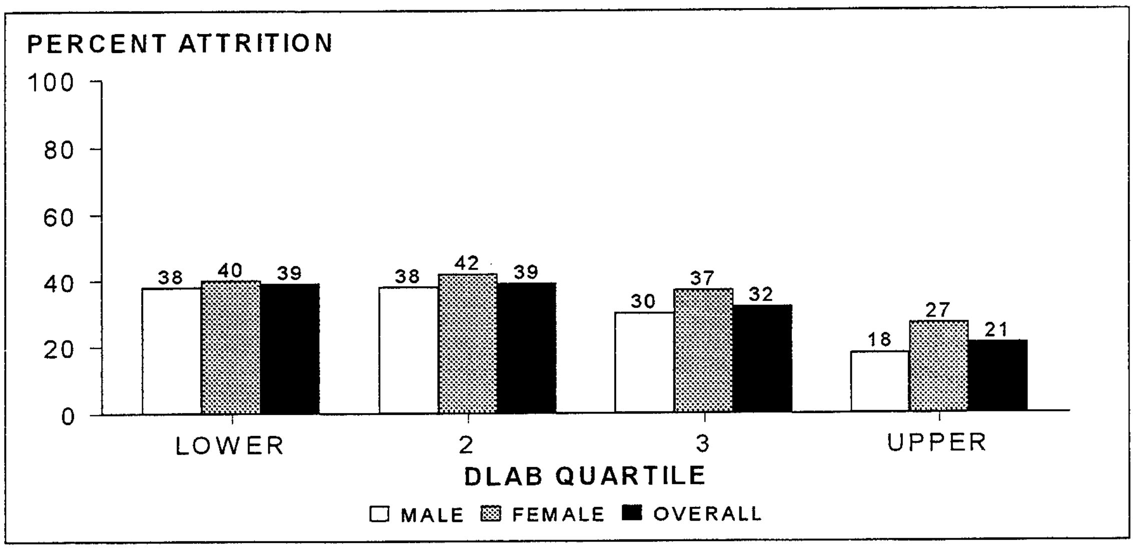

Figure 14. Percentage of attrition vs. DLAB.

greatest variance among all linear combinations of the predictor variables. The second principle component is the linear combination of predictor variables that has the greatest variance among all those linear combinations that are orthogonal to the first, and so on. The principle components are derived from an eigenvalue decomposition of the correlation matrix for the standardized variables, or the covariance matrix for the original variables. For variables that are measured on dissimilar scales it is important to perform principle components on standardized variables. Since ASVAB test scores are standardized, principle components on the original and standardized variables yield similar results. (Hamilton, 1992)

Let $x_1, x_2, \ldots, x_k$ represent the *n x 1* vectors of scores for each of the *k* tests, where *n* = 1,985 observations and *k* = 10 tests. The corresponding vectors of standardized test scores $z_1, z_2, \ldots, z_k$ are defined as :

$$z_j = s_j^{-1}(x_j - \bar{x}_j \cdot \mathbf{1}) \qquad \text{for } j = 1, \ldots, k \tag{2.1}$$

where $\bar{x}_j$ is the average over all observations for the j^{th} test, s_j is the standard deviation for the j^{th} test, and $\mathbf{1}$ represents the *n x 1* vector of ones required to make the vectors conformable. The first principle component of the correlation matrix is :

$$\alpha_1 z_1 + \alpha_2 z_2 + \ldots \alpha_k z_k \ , \tag{2.2}$$

where $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the first eigenvector of the correlation matrix and the $\alpha's$ are the loadings of each of the vectors of standardized variables. With subtest abbreviations as subscripts, values for $\alpha_{GS}, \alpha_{AR}, \alpha_{WK}, \alpha_{PC}, \alpha_{NO}, \alpha_{CS}, \alpha_{AS}, \alpha_{MK}, \alpha_{MC}, \alpha_{EI}$ respectively, are: (.34, .31, .34, .34, .31, .28, .28, .33, .31, .31). As shown in Figure 15, the first principle component accounts for approximately 68 percent of the variation in the ASVAB test scores.

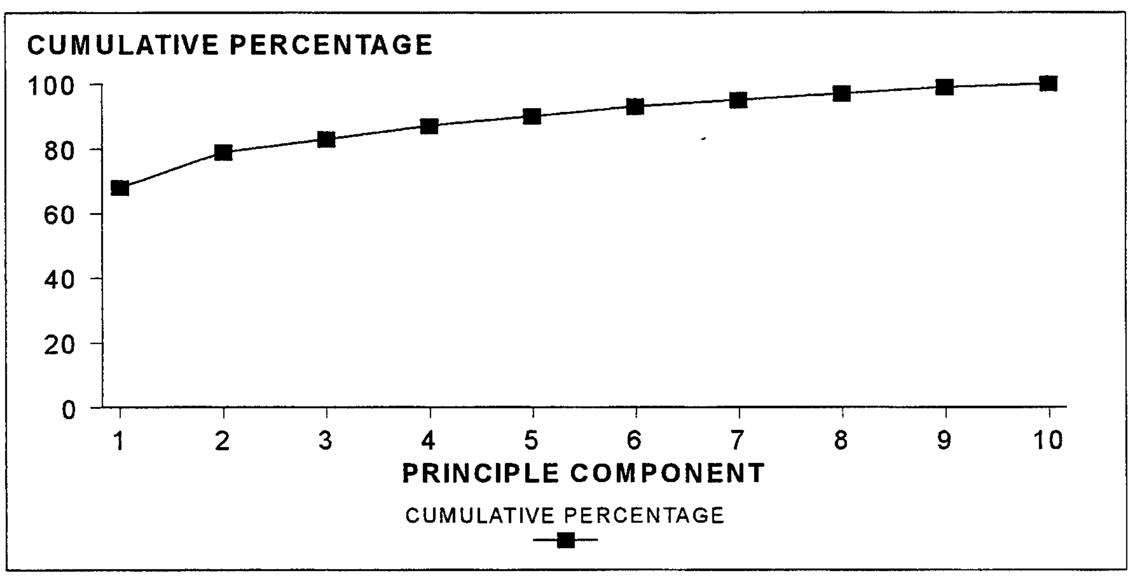

Figure 15. Cumulative percentage of variation in ASVAB test scores attributed to each principle component.

Equation (2.2) can be translated into the original test scores by replacing the standardized variables with the original variables, giving:

$$\frac{\alpha_1}{s_1}x_1 + \frac{\alpha_2}{s_2}x_2 + \ldots + \frac{\alpha_k}{s_k}x_k - \left(\sum_{j=1}^{k} \frac{\alpha_j}{s_j}\bar{x}_j\right) \cdot 1 \ . \qquad (2.3)$$

Thus, the first principle component corresponds to a weighted average of the original variables, where the weights are the loadings divided by the standard deviation of that variable. As shown in Figure 16, the loadings and standard deviations are about the same for each of the variables.

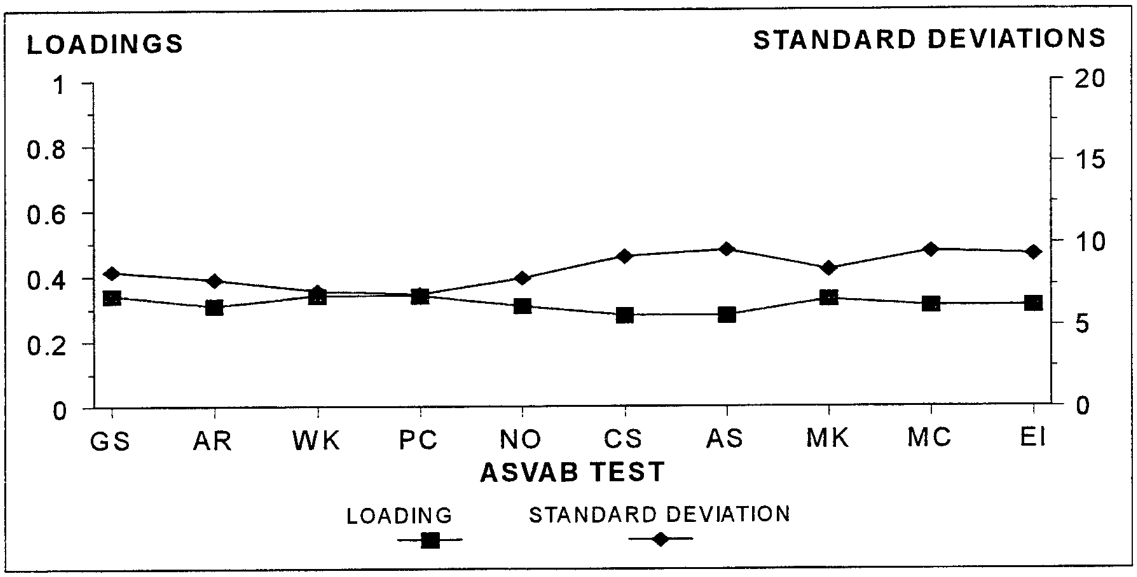

Figure 16. Loadings and standard deviations for each subtest in the first principle component of ASVAB scores.

The fact that the first principle component accounts for most of the variation in the test scores, and that the loadings for each of the factors in that principle component are about equal, means that an average of the test scores (weighting each test equally) accounts for the bulk of the variation in the ASVAB scores. Thus, a new variable, ASVABavg, was computed and is included in favor of the subtest scores. The net result of the principle component analysis is the reduction in the dimension of the ASVAB scores from ten to one.

The ASVAB and AFQT scores are other cases where there are a significant number of missing values. For these variables, approximately 30 percent of the observations have missing values. Approximately 40 percent of these missing values are attributed to the subjects being officers, because officers do not routinely take ASVAB tests. The remainder of the missing values for these variables are unexplained, but appear to be equally distributed among the other variables and have no other common attributes. As in the case of EDUYR, there is a concern over the handling of observations with missing values. Left uncorrected, this situation would lead to the exclusion of all officers and about 21% of enlisted observations.

To prevent the complete exclusion of observations with missing values for ASVABavg and AFQTavg, these variables are transformed from continuous to ordinal variables. Each observation is separated into its appropriate quartile, producing four categories. Then, the observations with missing values are placed into a fifth category. Thus, the variables ASVABqtiles and AFQTqtiles are included in favor of ASVABavg and AFQTavg. In this manner, observations with missing values for ASVABavg and AFQTavg can be included in the analysis across the entire range of predictors. As depicted in Figures 17 and 18, there is a generally decreasing percentage of attrition as test scores increase.

Variables indicating ASVAB and AFQT test versions (ASVFM and TESTV, respectively) are excluded, since these test scores are standardized and are therefore comparable without regard to test version.

Upon successful completion of study at the Institute, students are administered the Defense Language Proficiency Tests - Listening, Reading, and Speaking (DLPTL, DLPTR, DLPTS). Variables listing scores for the DLPTs are

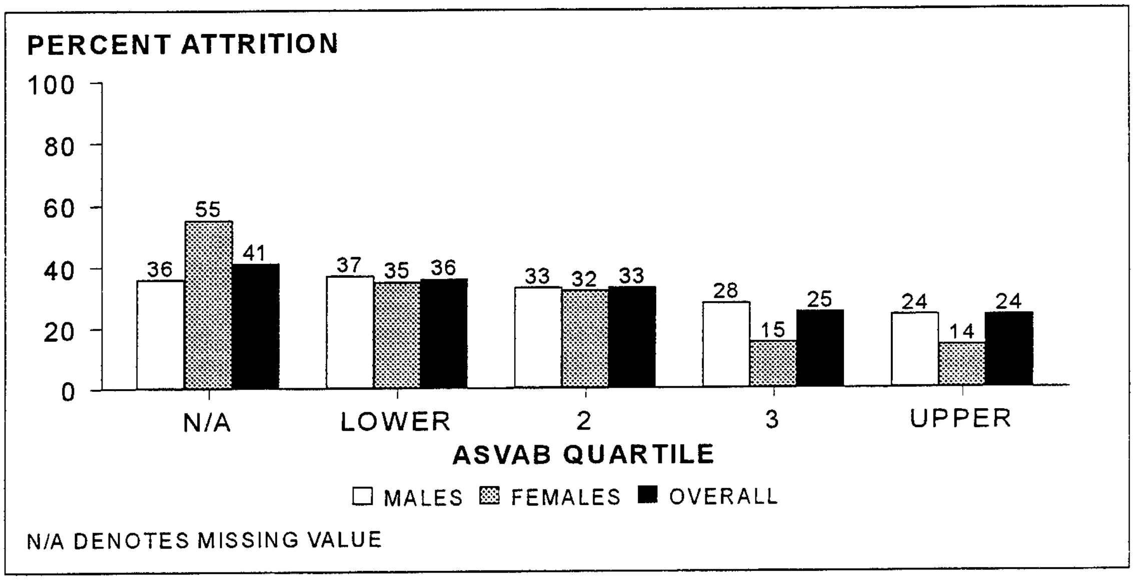

Figure 17. Percentage of attrition vs. ASVAB test scores.

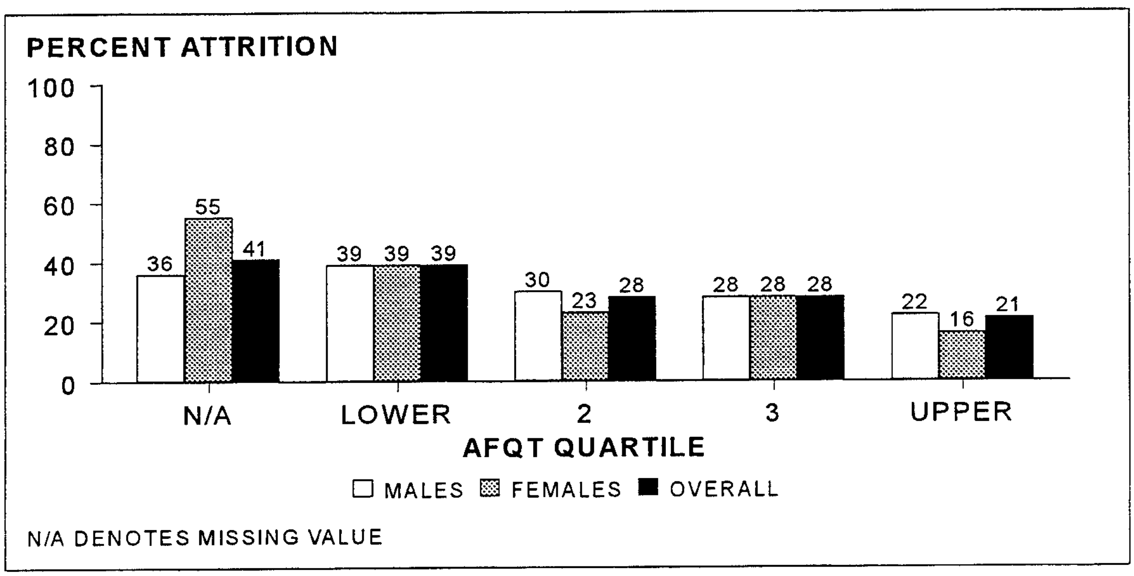

Figure 18. Percentage of attrition vs. AFQT test scores.

excluded as they are not available for students who attrite. Similarly, the variable listing a student's grade point average while at the Institute (GPA) is excluded since it is only recorded upon successful completion of the program.

After undergoing the preceding preliminary analysis, the data set includes the binary response variable GRAD/ATTR and the fifteen predictor variables listed in Table 2. These variables are used in modeling and further analysis of attrition. Development of a regression model of attrition is found in Chapter III.

VARIABLE	DESCRIPTION	TYPE	LEVELS
LANCAT	language difficulty category	ordinal	I, II, III, IV
DLAB	defense language aptitude battery test scores	continuous	N/A
YRSRV	years of military service	continuous	N/A
MARRY	marital status	nominal	married, single
MOTIV	level of motivation. self evaluated by student	ordinal	1. least motivated 2. . . 3. . . 4. . . 5. most motivated
PRILANCAT	prior language category. difficulty level of prior language, if any. compatible with LANCAT.	ordinal	I, II, III, IV
AGE	age as of 01JAN95	continuous	N/A
ETHNIC	ethnic category	nominal	0. unknown/none 1. white 2. black 3. hispanic 4. amer. indian/alaskan 5. asian/pacific islander 6. other
ASVABqtile	armed services vocational aptitude battery test score quartile	ordinal	0. missing value 1. lower quartile 2. second quartile 3. third quartile 4. upper quartile
AFQTqtile	armed forces qualification test (composite of ASVAB) quartile	ordinal	0. missing value 1. lower quartile 2. second quartile 3. third quartile 4. upper quartile
EDUYRgroup	highest year of education completed	nominal	N/A, HS, HS+
SERVICE	branch of service	nominal	USA, USAF, USN, USMC
PAYGRADE2	military paygrade	continuous	E1 = 1, ... , O6 = 20
OFF/ENL	officer/enlisted	nominal	officer, enlisted
SEX	gender	nominal	male, female

Table 2. Variables selected for use in analysis.

III. ANALYSIS

This chapter gives the details of the analysis performed on the data set and variables developed in Chapter II. The objective is to identify factors which have a significant impact on attrition ('significant' can mean either a positive or negative impact) with particular interest in those variables involving gender. The methodology involves developing a model of attrition, and further analyzing those variables which contribute significantly to the model.

A. THE MODEL

The data, prepared for analysis in Chapter II, include: a binary response variable (graduation/attrition) and a set of 15 predictor variables, which are a mixture of continuous and categorical variables (Chapter II, Table 2). Among the most common models considered appropriate for binary response variables are logit and probit. The logit model is used since the results from the two models are typically comparable, and the logit model is computationally easier to work with. (Collett, 1991)

Logistic regression fits binary response variables (Y) to a function of predictor variables $X_1, X_2, \ldots, X_p$ in such a way that E[Y], or equivalently, Pr(Y=1) is between 0 and 1. Specifically, it fits the logit of Pr(Y=1) as a linear function of the predictors $X_1, X_2, \ldots, X_p$ as follows:

$$\log\left(\frac{\Pr(Y=1)}{\Pr(Y=0)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p, \qquad (3.1)$$

or equivalently

$$\Pr(Y=1) = \frac{1}{1+\exp\{\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p\}}, \qquad (3.2)$$

where $\beta_0, \ldots, \beta_p$ are unknown parameters (Collett, 1991). In equations (3.1) and (3.2), Y =1 represents graduation and Y=0 represents attrition. Parameter estimates are obtained through the method of maximum likelihood, for which the logistic model has no closed form. Iterative numerical solutions are required; the most commonly used is Newton's method. Once the model is fitted, likelihood ratio tests are used to test the significance of the model as a whole and to eliminate variables which are redundant or do not have predictive ability. (Agresti, 1990)

B. ANALYSIS

Generally, there are two types of information which can be derived from any regression model. First, there is the ability of the model to *predict* changes in the response with respect to changes in predictor variables. Second, important insight into the question of interest may be obtained from the *structure* of the model itself; i.e., which predictors or combination of predictors seem to have a significant impact on the response. In the case of logistic regression, predictive ability is often limited (Hamilton, 1992). The typically low predictive power of logistic regression models is not a concern here, since the purpose of this study is not to predict who will attrite, but to compare attrition results between males and females.

1. Model Reduction

The first goal in arriving at a suitable model is to find a combination of predictors which capture the features of interest, yet is parsimonious. To reduce the number of predictors, a backwards elimination procedure is used. The first model is fit including all 15 main effects, and all of the two-way interaction terms between them (120 terms in all). Then, subsets of predictor variables which show the least significance are removed and the model is run again. This iterative procedure is continued until a satisfactory model is obtained with a balance of descriptive (not necessarily predictive) usefulness and simplicity.

Arrival at this satisfactory model is a matter of analyst judgment based on hypothesis testing. A comparison is made between the current (reduced) model and the one prior to it to test whether there is a significant difference between them. (Agresti, 1990)

The hypothesis test is performed as follows: let *Model(i)* represent the model under consideration in the i^{th} iteration of backwards elimination. Test the null hypothesis (*Ho*) that *Model(i)* is true versus the alternative hypothesis (*Ha*) that *Model(i-1)* is true. Note that under backwards elimination *Model(i)* contains fewer terms than *Model(i-1)*. Then the likelihood ratio test statistic (T) is two times the difference of the log likelihood under *Model(i)* and the log likelihood under *Model(i-1)*. The null distribution of T is approximately Chi-Squared with k degrees of freedom; k is the difference between the number of parameters in *Model(i)* and *Model(i-1)*. Large values of T indicate that the null hypothesis (Ho) should be rejected in favor of the alternative hypothesis (Ha); i.e., the model cannot be reduced by eliminating the variables chosen in the current iteration. Equivalently, if the p-value (the largest level of significance for which the test statistic causes rejection of *Ho*) is small, then *Ho* is rejected. If there is a significant difference between the models, then some or all of the removed effects should remain in the model. Main effects, regardless of significance, are left in the model if they are part of a significant interaction term. When no more effects can be removed from the model without a significant change, the current model is one which is as small (with respect to the number of predictor variables) as possible, and inferences can be made about the significance of the remaining predictors.

2. All Data

In all, 76 iterations were performed on the full data set. The final model includes 40 terms, of which 25 are significant (at a 0.10 level of significance). The uncertainty coefficient ($U = 0.2941$) indicates limited predictive power, as expected. The 'uncertainty coefficient' (U) is a statistic analogous to the familiar

R-squared, and its purpose is to describe the level of predictive utility in the model. It is computed as follows:

$$U = \frac{[-LogLikelihood(const\ model) - \{-LogLikelihood(fit\ model)\}]}{-LogLikelihood(const\ model)}, \qquad (3.3)$$

where the constant model is fit including only the intercept term. Table 3 lists significant terms in the final model, in order of decreasing significance. The p-values associated with the likelihood ratio test of the model excluding each variable, one at a time, are given in Table 3.

TERM	P-value	TERM	P-value
PAYGRADE2*AGE	0.0000	SERVICE*YRSRV	0.0069
MOTIV*ASVABqtiles	0.0000	MARRY*AFQTqtiles	0.0069
SERVICE*MOTIV	0.0000	DLAB*PRILANCAT	0.0074
SERVICE*ASVABqtiles	0.0001	LANCAT	0.0080
LANCAT*AFQTqtiles	0.0001	MARRY	0.0120
SERVICE*AFQTqtiles	0.0010	AGE*AFQTqtiles	0.0166
PRILANCAT*ASVABqtiles	0.0024	LANCAT*AGE	0.0258
YRSRV*AGE	0.0029	SERVICE*AGE	0.0265
LANCAT*PAYGRADE2	0.0032	DLAB*AFQTqtiles	0.0274
ETHNIC*ASVABqtiles	0.0035	PRILANCAT	0.0342
PAYGRADE2*EDUYRgroup	0.0039	SEX*SERVICE	0.0368
SERVICE	0.0055	PAYGRADE2*MARRY	0.0560
YRSRV*EDUYRgroup	0.0069	MOTIV	0.0938

Table 3. Significant terms in the final model, in order of decreasing significance.

Once the final model is developed, consisting of first order main effects and two-way interactions, further analysis is conducted to assure that the continuous main effects are of the proper form. Specifically, it is important to verify that the logit of the probability of graduation is linear in each continuous main effect and that transformations or re-parameterizations of the continuous main effects do not provide a better fit. Partial residuals are plotted against each continuous main effect. If the resulting plots are approximately linear, then higher order terms are not indicated. Partial residuals, PR_{ik} , are computed for each of the i = 1,..., 1985 observations and k = 1,.., 4 continuous main effects (PAYGRADE2, DLAB, AGE, YRSERV respectively), as follows:

$$PR_{ik} = \frac{Y_i - \hat{P}_i}{\hat{P}_i \bullet (1-\hat{P}_i)} + \hat{\beta}_k X_{ik} \quad , \qquad (3.4)$$

where:

Y_i = response (graduation/attrition) for the i^{th} observation,
$\hat{P}_i$ = estimated probability of graduation for the i^{th} observation,
$\hat{\beta}_k$ = parameter estimate for the k^{th} continuous main effect, and
X_{ik} = value of the k^{th} continuous main effect for the i^{th} observation.
(Collett, 1991).

From Figure 19, the plots of the partial residuals against DLAB, AGE, and YRSERV are quite linear, confirming that higher order terms are not required.

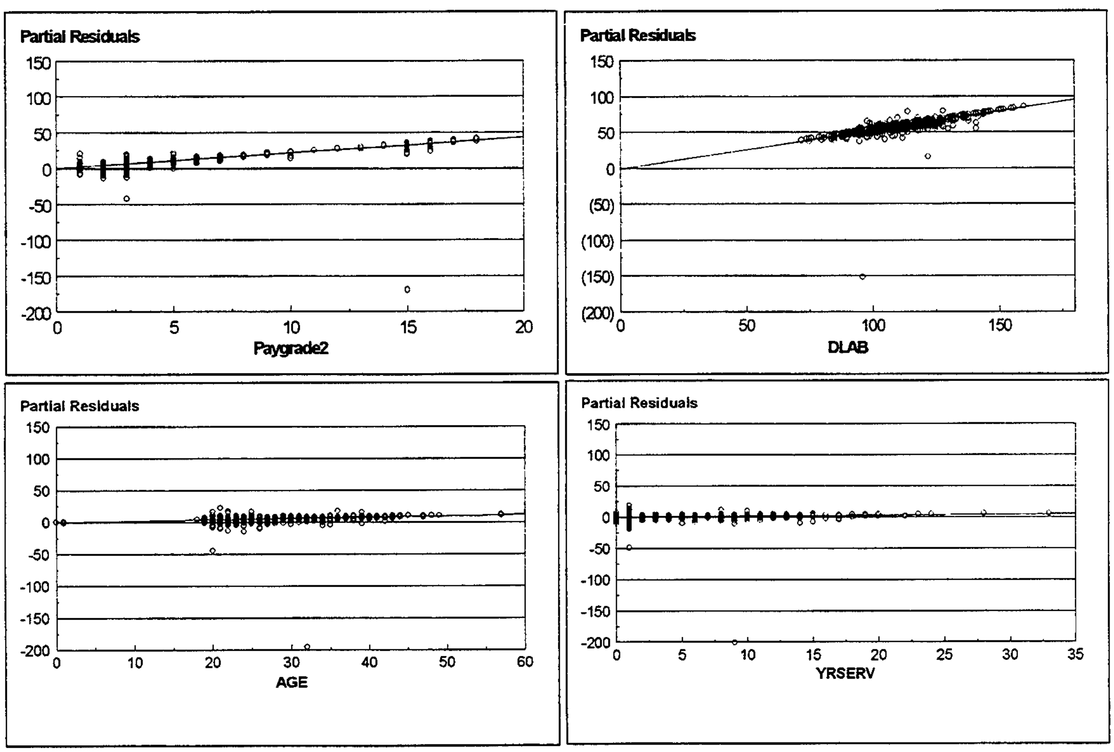

Figure 19. Partial residual plots for continuous main effects.

In the case of PAYGRADE2, there appears to be a slight non-linearity in the region of the lower paygrades. As a check, PAYGRADE2 is transformed into a categorical variable with one level for each paygrade. This transformation has no appreciable effect on the model, confirming that coding paygrade as the continuous variable PAYGRADE2 is adequate. Note that the slopes of the lines in Figure 19 are the parameter estimates for the respective variables, giving an indication of the relative impact of each of these variables on the model. A positive slope indicates a favorable impact on graduation as the values for these variables increase.

Analysis of the model structure will help to determine which variables have an impact on attrition. Figure 20 graphically depicts the complexity of the model given in Table 3. Of the main effects, 5 are significant:

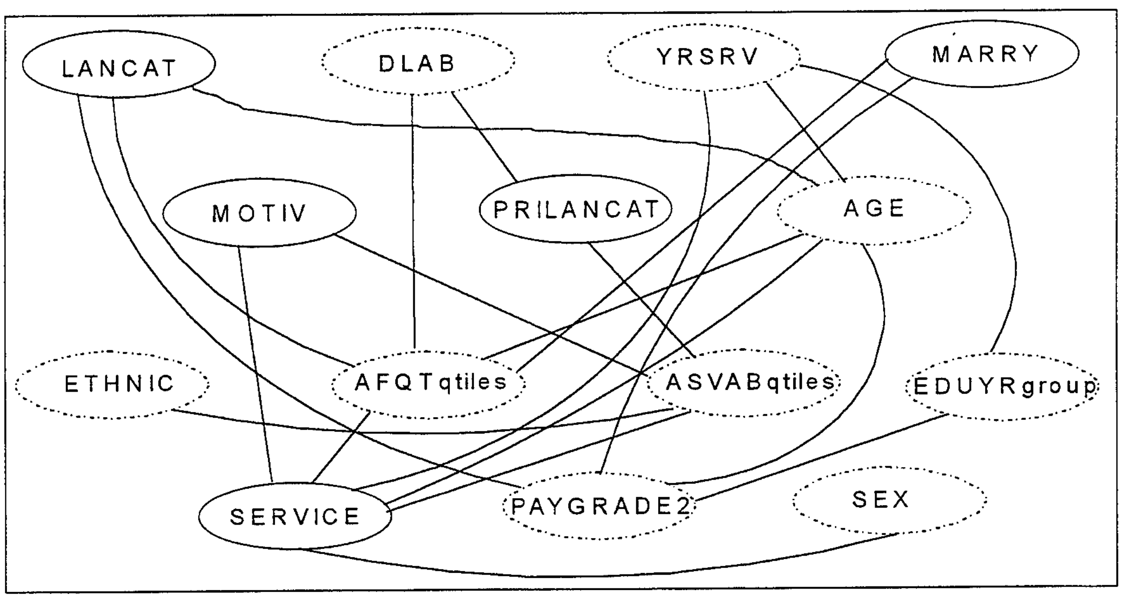

Figure 20. Graphical representation of variables in final model. Solid ellipses represent significant main effects, dashed ellipses represent non-significant main effects. Significant interaction terms are connected with lines.

LANCAT, PRILANCAT, MARRY, MOTIV, and SERVICE. Effects with a relatively high occurrence of interaction (4 or more) are: AGE, AFQTqtiles, ASVABqtiles, PAYGRADE2, and SERVICE. The predictor variable of interest, SEX, is not significant as a main effect, however its interaction with SERVICE is significant. The variable which appears to have the most impact, SERVICE, is significant as a main effect and is part of six significant interaction terms, most notable in this context is SEX*SERVICE. From Chapter II, Figure 10, we see that female Marines have a relatively high rate of attrition (60%). This suggests a possible explanation for the significance of the SEX*SERVICE interaction term.

3. Without USMC Data

A more detailed breakdown of the data is indicated. Specifically, the model is run again excluding all USMC observations to see if there is a change in the significance of the SEX*SERVICE interaction term. The iterative procedure described earlier in this chapter is used to reduce the model to the fewest possible number of predictors. Table 4 lists significant terms in the model run on data excluding USMC observations. Figure 21 graphically depicts the

TERM	P-value	TERM	P-value
PAYGRADE2*AGE	0.0000	YRSRV*EDUYRgroup	0.0075
SERVICE*MOTIVATION	0.0000	LANCAT	0.0092
SERVICE*ASVABqtiles	0.0002	DLAB*PRILANCAT	0.0095
SERVICE*AFQTqtiles	0.0003	LANCAT*AGE	0.0207
LANCAT*AFQTqtiles	0.0008	AGE*AFQTqtiles	0.0214
PAYGRADE2*EDUYRgroup	0.0021	ETHNIC*AFQTqtiles	0.0260
YRSRV*AGE	0.0024	DLAB*AFQTqtiles	0.0375
LANCAT*PAYGRADE2	0.0026	MARRY*AFQTqtiles	0.0411
PRILANCAT*ASVABqtiles	0.0026	MARRY	0.0412
SERVICE*AGE	0.0063	MOTIV*ASVABqtiles	0.0584
PRILANCAT	0.0065	SEX*SERVICE	0.0749
YRSRV	0.0072		

Table 4. Significant terms in the final model excluding USMC data, in decreasing order of significance.

information contained in Table 4. Excluding the USMC data reduces the complexity of the model slightly. The number of significant main effects is

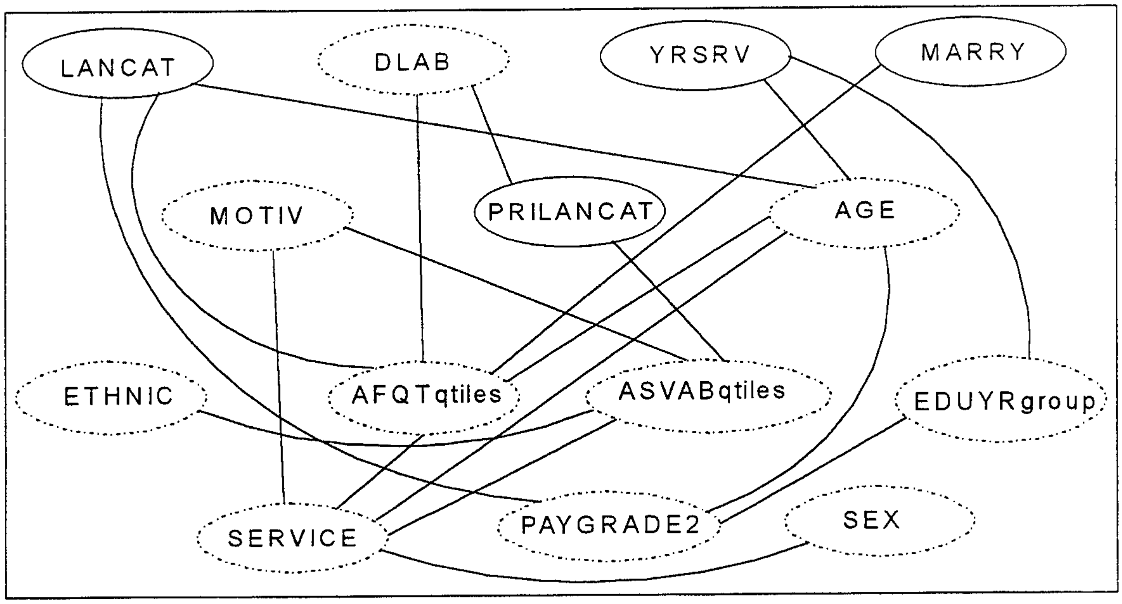

Figure 21. Graphical representation of variables in final model, excluding USMC data. Solid ellipses represent significant main effects, dashed ellipses represent non-significant main effects. Significant interaction terms are connected with lines.

reduced from 5 to 4 (MOTIV is no longer significant), while the number of significant interaction terms remains the same. Significant main effects include: LANCAT, YRSRV, MARRY, and PRILANCAT. Effects with a high frequency of interaction terms (4 or more) include: AGE, AFQTqtiles, ASVABqtiles, and SERVICE. The interaction term SEX*SERVICE is still significant (at a conservative level of significance of 0.10), although less so, with an increase in p-value from 0.0368 to 0.0749. The SEX*SERVICE interaction term was not affected greatly by controlling for USMC students, probably due to the relatively low weighting of USMC observations, which constitute only 5% (100 observations) of the data. Of all USMC observations, only 5% (5 observations) are female. In fact, the 60% (3 out of 5 observations) USMC female attrition rate has a standard error of 21%.

To further control for the effects of SERVICE, additional runs are performed on individual service groups. Computational problems arise when there are too many variables in a model, relative to the number of observations. Army and Air Force data are run individually, with 62% and 21% of the students, respectively. Navy and USMC data are not run individually, because they do not constitute a large enough proportion of the data to provide useful results (12% = 250 observations and 5% = 100 observations, respectively).

4. Army Data Only

Table 5 lists significant terms in the final model run for Army data, and Figure 22 provides a graphical representation of the information contained in Table 5. When compared to results from the model run including all data, there

TERM	P-value	TERM	P-value
LANCAT	0.0000	YRSRV*EDUYRgroup	0.0146
PAYGRADE2*AGE	0.0000	LANCAT*AGE	0.0162
LANCAT*AFQTqtiles	0.0001	PRILANCAT*ASVABqtiles	0.0212
PRILANCAT	0.0007	DLAB*AFQTqtiles	0.0215
YRSRV*AGE	0.0007	LANCAT*PAYGRADE2	0.0398
DLAB*PRILANCAT	0.0010	MARRY*AFQTqtiles	0.0488
YRSRV	0.0016	EDUYRgroup*AFQTqtiles	0.0520
LANCAT*DLAB	0.0030	AGE*AFQTqtiles	0.0576
PAYGRADE2*EDUYRgroup	0.0107	MARRY	0.0752

Table 5. Significant terms in the final model including only Army data, in decreasing order of significance.

is a reduction in the total number of significant terms from 25 to 18, with a reduction in the number of significant main effects from 5 to 4. Significant main effects include: LANCAT, YRSRV, MARRY, and PRILANCAT. SEX is not a significant predictor variable. Terms with a high frequency of interaction (4 or more) include: LANCAT, AGE, and AFQTqtiles. From Figure 22 we see a visible reduction in the overall complexity of the model for Army data only, as compared to the model run on all data.

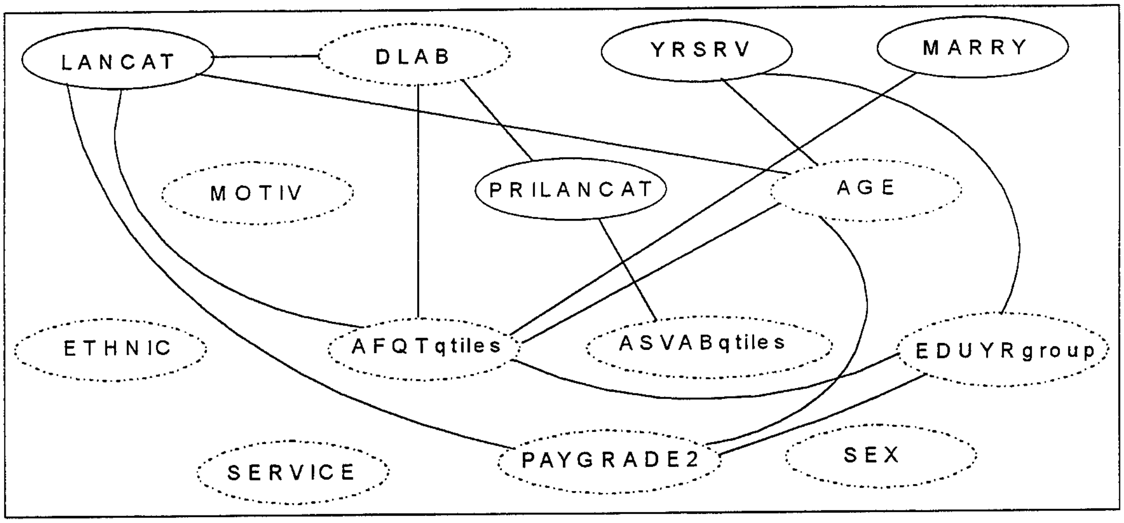

Figure 22. Graphical representation of variables in final model, including only Army data. Solid ellipses represent significant main effects, dashed ellipses represent non-significant main effects. Significant interaction terms are connected with lines.

5. Air Force Data Only

The next run was done on data including only Air Force students. Table 6 lists significant terms in this model. The smaller, less variable data set including

TERM	P-value	TERM	P-value
SEX	0.0180	AGE*AFQTqtiles	0.0545
MARRY	0.0296	LANCAT	0.0666
PAYGRADE2*MARRY	0.0389	PAYGRADE2	0.0721
MARRY*AFQTqtiles	0.0396	DLAB	0.0867
LANCAT*DLAB	0.0442		

Table 6. Significant terms in the final model including only Air Force data, in decreasing order of significance.

only Air Force students yields a much simpler model, resulting in only 9 significant terms, of which 5 are main effects: LANCAT, DLAB, MARRY, PAYGRADE2, and SEX. Of the 4 interaction terms, MARRY and AFQTqtiles are each involved in two. Most important is the fact that this group of observations results in the only occurrence of SEX as a significant main effect. The presence of SEX as a significant main effect for Air Force data probably explains the significance of the SEX*SERVICE interaction term for all data including Air Force observations. Figure 23 graphically depicts the relative simplicity of this model.

To confirm the suspicion that Air Force data causes the SEX*SERVICE interaction term to be significant, the model is run on the complete data set excluding Air Force observations. The SEX*SERVICE interaction term becomes highly insignificant, with a p-value of 0.3992.

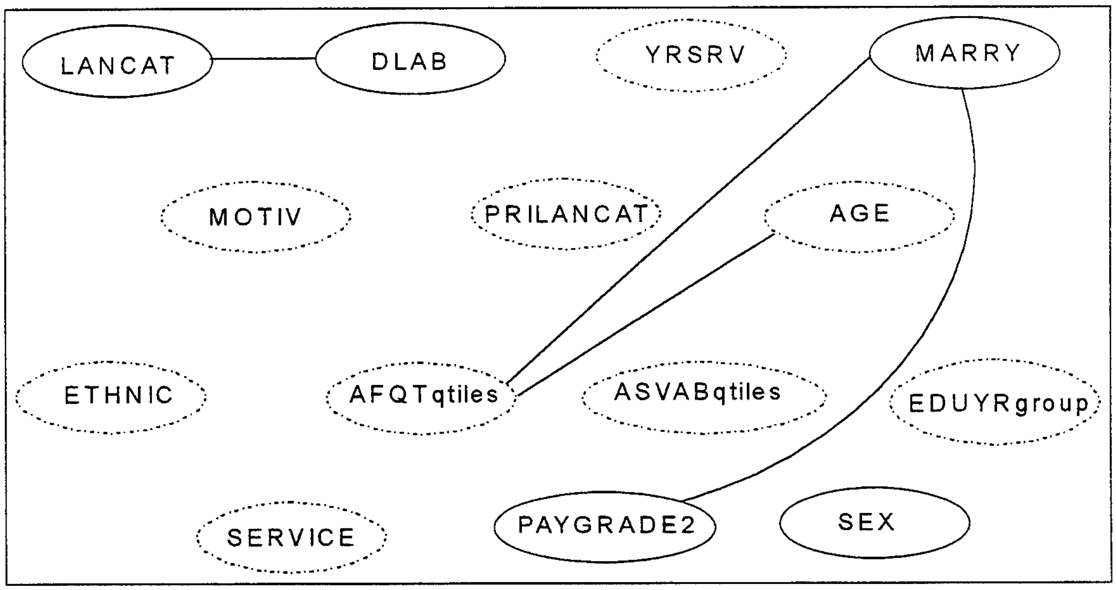

Figure 23. Graphical representation of variables in final model, including only Air Force data. Solid ellipses represent significant main effects, dashed ellipses represent non-significant main effects. Significant interaction terms are connected with lines.

To explore possible reasons why Air Force data might have this effect, a comparison is made between Air Force females and all other females in the data set. Key variables from each predictor block (demographic, language specific, and test scores) were chosen for comparison: SEX, AGE, YRSERV, PAYGRADE, LANCAT and DLAB. Females account for 35% of Air Force observations, compared to 24% for all other observations. There is no appreciable difference between groups for AGE, YRSERV, and DLAB. Air

Force females, however, are heavily weighted toward the more junior, 'high risk' paygrades, with 95% of Air Force females in paygrades E-3 and below, compared to 73% for all other females. 100% of Air Force females who attrited are from paygrades E-3 and below, compared to 84% for all other females. Language category distributions also differ, with 56% of Air Force females in the more difficult Category IV languages, compared to 45% for all other females. 60% of Air Force females who attrited are from Category IV languages, compared to 47% for all other females. These facts do not suggest that Air Force females are attriting more than their male counterparts due to their gender. In fact, for this model, equation (3.2) yields an estimated parameter value for the predictor variable SEX (with SEX coded as 0,1 for males and females, respectively) of approximately 0.50, with a standard error of 0.21. The positive value for this estimated parameter suggests the following; given exactly the same attributes (e.g., paygrade E-3 and below) a given female is likely to perform no worse than a male. For example, the probability of attrition for Air Force males in paygrade E-3 and below is 37%, compared to 36% for Air Force females. For paygrades E-4 and above, the probabilities are 16% and 0%, respectively. This is not inconsistent with the attrition statistics depicted in Figure 10, it merely underscores the fact that Air Force females tend to be in 'higher risk' paygrades.

6. Gender as Response Variable

An additional, less complex model is constructed to provide a different perspective on the problem. For this model, including only main effects, the roles of SEX and GRAD/ATTR are reversed, i.e., SEX is the response variable and GRAD/ATTR is used as a predictor. This is done to see if there is any change in the relationship between gender and attrition when viewed from this reverse 'angle'. If GRAD/ATTR is a significant predictor of SEX, then inferences can be made about the nature of the relationship between the variables. Table 7 lists all predictor variables in the model, with their associated p-values. There is

TERM	P-value	TERM	P-value
DLAB	0.0000	ETHNIC	0.0380
AFQTqtiles	0.0000	AGE	0.1027
ASVABqtiles	0.0000	LANCAT	0.1180
SERVICE	0.0000	GRD/ATTR	0.4496
EDUYR	0.0076	YRSRV	0.5121
PAYGRD	0.0248	PRILANCAT	0.5155
MARRY	0.0366	MOTIV	0.8516

Table 7. Terms in the final model with SEX as response, in decreasing order of significance.

a high degree of significance for variables related to test scores, and for branch of service. The p-value for GRD/ATTR is 0.4496, indicating that this variable is a highly insignificant predictor of gender.

Chapter IV contains conclusions based on the results of the analysis conducted in this chapter, along with recommendations for further study.

IV. RESULTS/CONCLUSIONS

This chapter summarizes results from Chapter III, and makes inferences about significant variables in the various models. Recommendations are made about areas which lend themselves to further study.

A. RESULTS

A final model is constructed for each of the five categories below. The models grow progressively simpler as the data groups become smaller and more homogeneous. Table 8 summarizes the results of the final model for each of the data groups included. Listed is whether the variable is significant as a main effect, and how many interaction terms it is involved in.

VARIABLE	ALL DATA	NO USMC	ARMY ONLY	AIR FORCE ONLY	SEX AS RESPONSE
LANCAT	Y/3	Y/3	Y/4	Y/1	N
DLAB	N/2	N/2	N/3	Y/1	Y
YRSRV	N/3	Y/2	Y/2	N/0	N
MARRY	Y/2	Y/1	Y/1	Y/2	Y
MOTIV	Y/2	N/2	N/0	N/0	N
PRILANCAT	Y/2	Y/2	Y/2	N/0	N
AGE	N/5	N/5	N/4	N/1	N
ETHNIC	N/1	N/1	N/0	N/0	Y
AFQTqtiles	N/5	N/5	N/5	N/2	Y
ASVABqtiles	N/4	N/4	N/1	N/0	Y
EDUYRgroup	N/2	N/2	N/3	N/0	Y
SERVICE	Y/6	N/5	N.A.	N.A.	Y
PAYGRADE2	N/4	N/3	N/3	Y/1	Y
SEX	N/1	N/1	N/0	Y/0	GRD/ATTR = N

Table 8. Variables in final model for each data group. Listed is significance as main effect/number of interaction terms the variable is involved in.

For the model including all of the data, there are 5 significant main effects. Variable blocks with the highest frequency of significant variables, either as main effects or interaction, are: demographics (5), language specific variables (2), and test scores (2). Service branch is the single most involved effect; it is significant as a main effect and is involved in 6 interaction terms. The predictor of interest, SEX, is not significant as a main effect, but its interaction with SERVICE is a significant effect (p-value = .0368).

To control for apparent anomalies in the attrition statistics for female Marines, the data are broken into smaller groups. The model is run on all data, excluding USMC observations, to see if the interaction term SEX*SERVICE remains significant. Controlling for the USMC data does not eliminate the interaction of SEX*SERVICE as a significant effect, although its p-value is increased from 0.0368 to 0.0749. Although removing USMC observations removes SERVICE as a significant main effect, it is still significant in several interactions. The model is not very sensitive to the exclusion of USMC data due to the small number (5) of female Marines in the data set.

To further investigate the effects of branch of service on attrition, additional runs are made on the Army and Air Force data seperately. There are too few observations for the other services (Navy and USMC) to allow fitting the model with all of the predictor variables.

For the model run on Army data only, there are 4 significant main effects. Variable blocks with the highest degree of involvement in significant effects include: demographics (3), language specific variables (2), and test scores (1). The predictor of interest, SEX, is not significant as a main effect or interaction term.

For the model run on USAF data only, there are 5 significant main effects. This is the only data group in which SEX is a significant main effect. The presence of SEX as a significant main effect for the Air Force data leads to the conclusion that the Air Force observations cause the significance of the SEX*SERVICE interaction term in models including Air Force data. A model run on all data, excluding Air Force observations, supports this conclusion since the

SEX*SERVICE interaction term becomes highly insignificant. Further analysis reveals attributes in which Air Force females differ from other females among the entire data set. Females account for 35% of all Air Force data, compared to 24% for the other services as a whole. Other areas in which Air Force females differ are language category and paygrade. 56% of Air Force females are in the most difficult language category (IV) compared to 45% for all other females. Also, 95% of Air Force females are in the 'higher risk' paygrades of E-3 and below, compared to 73% for all other females.

B. CONCLUSIONS

In summary, gender is a significant main effect for the model run on Air Force subjects only, and it is a significant interaction term for the full data set and the data excluding USMC observations. A model run on all data, excluding Air Force observations, supports the conclusion that the Air Force subjects cause the significance of the SEX*SERVICE interaction term in the other models.

This study indicates that Air Force females do not attrite more frequently than their male counterparts due to their gender; in fact, compared to Air Force males with identical attributes (e.g., the same paygrade group) Air Force females have similar (or smaller) attrition rates. The higher overall attrition rate for Air Force females is mostly due to their relatively high proportions in lower paygrades and more difficult language categories.

With the exception of the model in which SEX is the response, the language specific variables, LANCAT and PRILANCAT, consistently outperform other variable blocks, followed closely by demographic variables and assorted test scores. The significance of the block of demographic variables is consistent with the findings of the Language Skill Change Project referenced in Chapter I.

A final conclusion is that higher attrition rates for females do not appear to be attributable to their gender. Instead, particularly in the case of Air Force females (the group having the largest gender impact on the attrition model), the comparatively higher attrition rates are considered to be a function of relatively high proportions of females in 'higher risk' groups such as junior paygrades and more difficult language categories.

C. RECOMMENDATIONS FOR FURTHER STUDY

There are two areas which lend themselves to further study. First, the apparent impact of gender on attrition for Air Force students suggests that a more in depth analysis of Air Force students be conducted to further explore the causes for the significant relationship between gender and attrition for these students.

Second, a more detailed exploration of why students fail to graduate is indicated. Specifically, there appears to be an imbalance in these reasons for males and females. From Chapter I, recall that females attrite overall at a higher rate than males. However, attrition for academic reasons is much higher for males. 'Reason Out' data, as it is currently collected at DLIFLC, is broken into the following categories: Currently Enrolled, Academic, Physical Fitness, Lack of Effort, Over Weight, Medical, Discipline, Unit Recall, Security Clearance, and Other.

Excluding the Currently Enrolled and Academic categories, there is a relatively high use of the 'Other' category (approximately 15% overall). This appears to be at the expense of the remaining categories, suggesting a possible overuse of the 'Other' category. Overuse of the 'Other' category may result in the loss of information as to the true reason for some student losses. It would be useful to determine if this is in fact the case, and to correct the category assignment procedures, if necessary. This measure would facilitate a further analysis of the various reasons behind student attrition.

LIST OF REFERENCES

Agresti, A., *Categorical Data Analysis*, John Wiley & Sons, 1990.

Begley, S., "Grey Matters", *Newsweek*, 27 March 1995.

Collett, D., *Modelling Binary Data*, Chapman & Hall, 1991.

Directorate for Academic Administration, *DLIFLC Program Summary*, 1995.

Dove, M., Defense Manpower Data Center, telephone interview, July, 1996.

Hamilton, L., *Regression With Graphics*, Duxbury Press, 1991.

O'Mara, F., *Learning Skills Change Project Executive Summary*, DLIFLC, 1994.

O'Mara, F., *Learning Skills Change Project Report II*, DLIFLC, 1994.

Rice, J., *Army Linguist Personnel Study*, DLIFLC, 1975.

Shaw, V., Christie, E., *DLIFLC-DMDC Student Data Base Documantation*, DLIFLC, 1994.

INITIAL DISTRIBUTION LIST

		Number of Copies
1.	Defense Technical Information Center 8725 John J. Kingman Road., Ste 0944 Ft. Belvoir, VA 22060-6218	2
2.	Dudley Knox Library Naval Postgraduate School 411 Dyer Rd. Monterey, CA 93943-5101	2
3.	Lyn Whitaker Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5101	2
4.	Sea Control Squadron Fourty One PO BOX 357098, N.A.S. North Island San Diego, CA 92135 - 7098 ATTN: LT George Arthur, USN	2
5.	Stephen Payne Command Historian Defense Language Institute Foreign Language Center Presidio of Monterey, CA 93944-5006	2
6.	John Lett Director, Research and Analysis Division Defense Language Institute Foreign Language Center Presidio of Monterey, CA 93944-5006	3
7.	AISO Library Defense Language Institute Foreign Language Center Presidio of Monterey, CA 93944-5006	1